Reuse Enablement System (RES) Trade Study Report

Prepared by:
NASA Earth Science Data Systems –
Software Reuse Working Group

November 17, 2005

Earth Science Data Systems Software Reuse Working Group

Editor:

James Marshall (Innovim / NASA Goddard Space Flight Center)

Contributing Working Group Members:

Nancy Casey (Science Systems and Applications, Inc. / NASA Goddard Space Flight Center)

Victor Delnore * (NASA Langley Research Center)

Robert R. Downs (Columbia University / NASA Socioeconomic Data and Applications Center)

Stefan Falke (Washington University in St. Louis)

Ryan Gerard (Innovim / NASA Goddard Space Flight Center)

David Giles (Science Systems and Applications, Inc. / NASA Goddard Space Flight Center)

Tommy Jasmin (University of Wisconsin-Madison, Space Science and Engineering Center)

Steve Olding (Everware / NASA Goddard Space Flight Center)

Shahin Samadi (Innovim / NASA Goddard Space Flight Center)

Ross Swick (National Snow and Ice Data Center, University of Colorado – Boulder)

Bill Teng (Science Systems and Applications, Inc. / NASA Goddard Space Flight Center)

Robert Wolfe * (NASA Goddard Space Flight Center)

* Co-chair

Working Group Participants:

Steve Ackerman (University of Wisconsin-Madison)

Bruce Barkstrom (NASA Langley Research Center)

Clyde Brown (Science Applications International Corporation / NASA Langley Research Center)

Howard Burrows (Autonomous Undersea Systems Institute / National Science Digital Library)

Bradford Castalia (University of Arizona)

Yonsook Enloe (SGT Inc. / NASA Goddard Space Flight Center)

Albert Fleig (Science Systems and Applications, Inc. / NASA Goddard Space Flight Center)

Michael Folk (National Center for Supercomputing Applications)

Claude Freaner (NASA Headquarters / NASA Jet Propulsion Laboratory)

Emily Greene (Raytheon Company)

Watson Gregg (NASA Goddard Space Flight Center)

Bob Harberts (Global Science & Technology, Inc.)

Rudolf Husar (Washington University in St. Louis)

Steven Kempler (NASA Goddard Space Flight Center)

Scott Lewicki (NASA Jet Propulsion Laboratory)

Michael Little (NASA Langley Research Center)

Christopher Lynnes (NASA Goddard Space Fight Center)

Edward Masuoka (NASA Goddard Space Flight Center)

Patrick Moran (NASA Ames Research Center)

Ruth Neilan (NASA Jet Propulsion Laboratory)

Susan Paradise (NASA Jet Propulsion Laboratory)

Glen Schuster (US Satellite Laboratory / Signals of Spring)

Ilya Slutsker (Science Systems and Applications, Inc. / NASA Goddard Space Flight Center)

Fred Watson (California State University, Monterey Bay)

Keith Wichmann (Global Science & Technology, Inc.)

Thomas Yunck (NASA Jet Propulsion Laboratory)

Feng Zhong (Global Science & Technology, Inc.)

Table of Contents

Executive Summary

As the number of Earth observing instruments and captured data volume increase, so do the complexity and costs associated with software development in support of data transformation, analysis, processing, management, and end-product implementation. Software development costs can be high and the time needed to develop new applications can be considerable. The Earth Observing missions have aided in amplifying knowledge of the Earth system by generating many useful scientific data. To maximize the use of these data, the Earth science community must be able to spend less time, money, and effort on software development and more on scientific work. Reusing software, including open source software, has many benefits such as increased productivity, reduced schedule, and improved quality. However, realizing these benefits for Earth science has been challenging.

Our survey of the Earth science community has indicated that lack of a centralized domain-specific software repository or catalog system addressing the needs of the Earth science community is a major barrier to software reuse within the community. The Software Reuse Working Group was tasked by NASA HQ to "undertake a trade study to understand the roles of NASA Open Source Agreement web site (http://opensource.arc.nasa.gov), the role of GCMD in cataloging data services, and any other related NASA software sites" as a potential platform to enable software reuse for the Earth science community. These two sites and five other NASA sites were reviewed. A variety of non-NASA sites performing similar roles was also reviewed. We were also asked to keep in mind that the Earth and space science communities have similar reuse needs.

This report provides a list of the key requirements, identified by the Working Group, that are necessary for a software repository or catalog system to meet the needs of the community for storing and locating reusable software assets. It then covers the findings of our evaluation of existing systems based on these requirements.

The results of our study show that none of the evaluated repository or catalog systems can satisfy the needs of the community. The shortcomings typically fall under the following general areas: they do not meet enough of the critical functional requirements, they do not focus on the Earth science domain, they do not target software developers as their primary audience, and they do not provide the types of small-sized assets that are most useful to the community of Earth science software developers. In particular, the need to support multiple licensing mechanisms for reusable assets and the need to restrict access to some assets to meet export control and NASA release requirements provide particular obstacles to using most of the systems evaluated. The GCMD provides information about a large number of data sets in many areas of Earth science, but does not satisfy the functional requirements, does not target software developers, and does not provide the type of assets that are most useful to software developers. The GSFC and Ames open source sites are useful for promoting NASA's open source software projects, but have the same problems as the GCMD and add restrictions on license and organization – only open source products produced by NASA are available. Furthermore, these sites have no underlying catalog/repository functionality. Assets are simply listed as links on a web page.

The findings of our trade study demonstrate that currently there is no repository or catalog system that is capable of addressing the needs of our community. This confirms the

survey results where community members expressed the opinion that lack of such a system was a major barrier to software reuse. We further conclude that none of them would be close enough to meet our needs without considerable re-engineering and development. As a result of this study, we recommend that NASA provide the necessary support for a reuse enablement system dedicated to the Earth science community that could be expanded to include the space science community. We recognize that there is a potential for collaboration between a new reuse enablement system and current catalogs and repositories such as the GCMD. Based on our existing knowledge, we expect that currently available cataloging and repository technologies would be able to fulfill a significant part of the required functionality of a reuse enablement system, but a more detailed architecture study must be performed to determine the most expeditious and cost-effective solution for such a system.

Background

To address the technical issues required to enable and facilitate reuse of software assets within NASA's Earth Science Enterprise (ESE), the NASA Earth Science Software Reuse Working Group was created as part of the Earth Science Data System (ESDS) Working Group. This was the result of one of the recommendations from the NASA HQ-commissioned Strategic Evolution of ESE Data Systems (SEEDS) Study; the SEEDS activity became the ESDS Working Group activity. The Software Reuse Working Group was chartered to oversee the process that will maximize the reuse potential of such software components in order to: (1) drive down the cost and time of system development and reduce/eliminate unnecessary duplication of effort; (2) increase flexibility and responsiveness relative to Earth science community needs and technological opportunities; and (3) increase effective and accountable community participation.

The Working Group recommends and supports activities that help increase awareness of available software components, increase awareness of the value of reuse, provide needed processes and mechanisms, disseminate successful reuse strategies, and address related intellectual property and policy issues. In the process of fostering greater software reuse across the Earth science systems, the Working Group is considering a wide variety of approaches to help meet differing needs and priorities. One such approach has been the creation of the Software Reuse Working Group portal web site at http://softwarereuse.nasa.gov/. It contains information on reusable assets, resources such as events and publications, open source software in general, and funding opportunities as well as some information about the Working Group's projects. Thus, it provides users with a central location for finding information about software reuse.

The goal of the software reuse project is to encourage software developers to make use of existing software assets, including open source software, to provide them with a convenient way to locate and obtain such assets, and to encourage them to develop products for reuse by others. The process of creating a new software product by reusing existing components can be likened to the building of a house. The consumers will be able to buy a completed house, but it is the builders who create the house from a variety of pre-fabricated components such as the frame, windows, and plumbing. By using tools, parts, and methods that have been tested over time and are known to work well, it becomes easier and more efficient for them to build the house. Likewise, if software

developers can make use of existing software components, it will be easier and more efficient for them to create new products.

Software released under an open source license is publicly available and other software developers can read, modify, and redistribute the source code. The Working Group has recommended greater use of open source licensing as an important enabler for software reuse. The simpler licensing mechanism of open source, compared to traditional software licensing, eliminates a significant barrier to code sharing and thus helps to encourage and promote reuse. However, open source licensing is not appropriate for all types of software and traditionally licensed software can still be reused. Therefore, an effective reuse program has to accommodate both open source and non-open source software.

To facilitate the software reuse process, developers need to be able to easily locate and evaluate the available reusable artifacts. These were identified as important factors in a survey (OMB #2700-0117) conducted by the Working Group to determine the reuse practices of the Earth science community. See Appendix A for additional information about the survey.[1] The results showed that when people did not reuse software, the primary reasons were because they did not know where to look and they did not know such reusable software existed. In addition, the survey revealed that a catalog or repository for reusable artifacts is the best means of increasing software reuse within the Earth science community. For this reason, the reusable artifacts should be classified and made available through an appropriate reuse enablement system (e.g., libraries, catalogs, repositories) that can facilitate searching and indexing. These systems are an essential ingredient in transforming ad-hoc reuse (which is largely dependent on personal knowledge and word of mouth dissemination of information about the availability of reusable artifacts) to systematic reuse as an integral part of the software development process.

To achieve the above goal, the Working Group was tasked to research and evaluate existing software catalog and repository systems within NASA, specifically the GCMD and the NASA Open Source Agreement site, as possible alternatives to: (1) hosting software assets for the Earth science community and/or (2) developing an Earth science Reuse Enablement System by using existing enablement system reusable infrastructure software components. See Appendix B for the Working Group's original recommendation and NASA HQ's response. These sites as well as other NASA sites and a variety of non-NASA sites performing similar roles were examined and reviewed. The focus of the study is on Earth science related software repositories and catalogs, but recognizing that the scope may be expanded in the future to include space science, a few astronomy related sites are included in the review. Also, some of what has been developed for the space science community may be useful to the Earth science community and vice versa. The requirements against which the sites were reviewed are described in the following section.

---

[1] See also the Proceedings of the 2004 IEEE International Geoscience and Remote Sensing Symposium, vol. 3, pp. 2196-2199; "Strategies for Enabling Software Reuse within the Earth Science Community" by Samadi et al. for preliminary results from an earlier, almost identical survey.

Earth Science Reuse Enablement System Requirements

The primary function of a Reuse Enablement System (RES) is to facilitate the distribution and reuse of software development artifacts across the Earth science community, with possible extension to space science. The reusable artifacts supported by the system will include software components and other digital artifacts used in the software development process. It must also have the ability to function as both a catalog and a repository, the main difference between the two being that a catalog stores links to artifacts while a repository stores the actual artifacts themselves. By being able to perform both tasks, the RES provides more options to the user and is able to go beyond what either a catalog or a repository can do alone.

The primary users for the Reuse Enablement System are NASA-funded software developers within the Earth science community who create products. The other category of user is researchers and scientists in various organizations.   Many of these are researchers from the Earth science community who may be involved with NASA projects. Others are academic scientists or members of research communities. In some cases, the users are also asset providers, implementing software assets and delivering them to the RES for dissemination.

> The Software Reuse Working Group conducted a workshop to identify the functional requirements needed for a software Reuse Enablement System (RES) supporting the Earth science community. Several members of the Working Group participated in this workshop and helped draft the initial set of requirements. Over a period of several months, these requirements were refined through weekly and monthly telecons and finalized during a review at the October 2004 Data Systems Working Group meeting. The result of this work identified a number of requirements in the following areas: general, search, user registration, asset usage, asset submission, content management, and system administration. These are described in detail in Appendix C.

The general requirements include the kind of features that all systems are expected to exhibit, such as supporting remote access through standard Internet browsers and allowing administrators to generate reports including metrics.   It also includes non-functional requirements such as the system being in the Earth science domain, providing the types of assets that are most useful to software developers, and providing a method for appraising the submitted assets for quality control purposes. The focus on the Earth science domain was considered important because it enables the repository to have an asset classification system specific to the needs of the target audience and Earth science assets would not be obscured by large numbers of non-relevant artifacts. The functional requirements can be viewed in the larger categories as noted above (see Appendix C) or in smaller categories based on the specific functions they provide, the use cases for the system.[2]   For example, one obvious required function is that the system shall allow users to acquire an asset from the system.   Using the latter viewpoint, the functional requirements are described in detail in the following sections.   However, all of these requirements are not fixed and during the implementation stage, they may be revised based on cost/benefit analysis, available resources, etc.

---

[2] http://www.sciencedatasystems.org/seeds/wg/reuse/ -- Working Group Documents, Support and Enablement, Reuse Enablement System Use Cases

Register User

The system shall allow a user to register with the system. The user will have the choice to register as either a consumer, who will only download content, or as a provider, who will upload content as well as download it. The user will then select a user name and password for use on the system, and provide some identification and contact details. Once the registration has been successfully submitted, an administrator will review the registration request and decide whether to accept or reject it. The user will be notified of the result and accepted users will be successfully registered in the system. All registered users will be able to update the information in their profiles, but major changes such as requesting a change from consumer to provider may be subject to administrator approval. If it is necessary to restrict access to some assets, for example due to export restrictions, users can be divided into different groups and permission to access those assets will only be granted to the appropriate group(s). Administrators will have the ability to manage user accounts and passwords.

Registration will not be required simply to view and browse through the system to see what is available. However, it will be required to submit assets and may be required for downloading assets. This provides some degree of security since the administrators will have information on all users who are downloading or uploading content and know everyone who has this level of access to the system. It is also a means of quality control since all assets that are submitted will have a known source or provider who can be contacted if necessary. Related quality control measures include only allowing upload access to users from trusted organizations (e.g., NASA, other government agencies, colleges, universities) and requiring providers to submit assets under their real names, not pseudonyms. The former grants access based on the employer's reputation for quality output while the latter helps ensure quality by making the provider stand behind his/her submitted products.

Contribute/Update Assets

Registered providers shall have the ability to upload assets to the system. This feature allows providers to submit assets for inclusion in the repository or catalog. The user shall follow a process to provide adequate information about the asset such as its name, a description of it, links to supporting documents, the language the software was written in, the platforms on which it runs, a Technology Readiness Level (TRL) or other assessment of maturity, in addition to the asset itself and/or a link to the asset. Other users will be able to search for assets based on this information. Submitted assets will not be available to the community immediately; a content manager must check and approve the asset before it will be made available. Providers will also be able to update their submitted assets, but major changes such as uploading a new version that changes the overall functionality of the software may require the approval of a content manager.

In order to meet the needs of a growing and evolving community, a repository or catalog must be able to grow and evolve along with it. Allowing registered users of the system to submit assets to it is the best way to make sure that the community's needs are being met. Users will be providing assets that they themselves use and may be of use to others. Requiring the approval of a content manager before assets are made publicly available is primarily a way of confirming that the assets are suitable for inclusion in the system; e.g., that their subject matter falls within the domain of the system.

System Feedback

The system shall allow users to provide feedback about their experiences with the system to its administrators and managers. The user will be able to follow a link on the site that leads to a feedback page where a form may be completed and submitted. Bug reports, performance problems, and suggestions for enhancements are examples of the kind of system feedback users may supply. This feature may be available only to registered users.

Users can assist the system managers in keeping the system running properly and efficiently by providing feedback about its performance and usability. This is useful for maintaining the system as it alerts the managers to problems as soon as they are discovered by users, allowing them to be fixed promptly. The feedback can also help helpful in improving the system content and for planning future system development. Feedback may be accepted only from registered users in order to avoid potential abuse or misuse of anonymous feedback.

Automatic Notifications

The system shall offer all registered users the option to receive automatic notifications about updates to the system (e.g., via e-mail). This will be handled through subscriptions that allow users to select which types of events they wish to be notified. The upload of a new version of an asset and the change of an asset description are examples of the types of events that may be available for subscription. When such an event occurs, the system will automatically notify all users who are subscribed to that event.

Automatic notifications allow users to stay informed about system or asset activities that interest them without having to look for such activity themselves. For example, new versions of software generally fix problems in older versions while improving existing features and/or introducing new ones. Therefore, users often want to run the latest version. Providers can use the subscriptions to know when a new comment about their asset has been added to the system so that they may answer users' questions and assist them if they have difficulty using the asset. By having the system automatically notify users of updates when they are available, it saves the users the time and effort of performing frequent manual requests.

Discovering Assets

The system shall provide the following ways for users to see what is available and allow them to find a product that meets their needs: providing a complete list of all assets for the user to browse, classifying the assets to form a hierarchy and allowing the user to navigate through this structure, and allowing the user to perform a search for the desired asset. The ability to search on keywords such as type of asset or language, the information that content providers give when submitting an asset, is an important feature as it can narrow down the search results, enabling users to find suitable products more quickly. In addition, the user shall be able to order the search results in a variety of ways such as by category or rating.

Before a user can download an asset to use, the desired product must be located. This can be done in a variety of ways, but a search option is generally the least restrictive. If the user wishes to place additional limits on the search, the product information given by

the asset providers can be used to filter the results. Once a suitable asset is found, by any means, the user can download it to use.

Register Asset Usage

The system shall allow users of the system to register the usage of an asset. This enables users to indicate that they are actively using the asset, something that downloading alone does not indicate. The registration can be used to generate statistics about the reuse of the asset and to keep users informed about bug fixes, updates, enhancements, and other similar events for the asset.

It is important to know when an asset is actively being used in order to keep its users informed about events related to it and to generate accurate statistics on reuse. Reuse statistics are intended to be derived from people who are actively using the product, rather than just sampling it. Not everyone who downloads an asset will become an active user of it, so statistics should not be based purely on the number of downloads. Allowing users to register their usage of assets provides an explicit indication of active usage and thus is a better measure for statistics.

Asset Review

The system shall allow users to provide feedback about their experiences with assets in the system. After selecting one of the assets in the system, the user will be able to follow a link to a review page. Here the user will be able to submit a review of and a rating (e.g., on a scale of 1-5) for the asset as well as see the reviews and ratings of other users. Also, individual ratings will be averaged together to create an average rating for the asset. Installation and usage experiences (good and bad), bug reports, performance problems, and suggestions for enhancements are examples of the kind of information users may supply in their review. Asset providers may also be able to respond to reviews through this feature, for example, to provide advice to users about issues raised in their reviews. Submitting reviews may be available only to registered users, but all users will be able to see reviews. Submitted comments and reviews are subject to the approval of a content manager.

The ability to provide asset reviews is an important feature since it is helpful to everyone who uses or runs the system. The reviews that users provide will highlight any good and bad points about the assets, enabling other users to make a more informed choice about what assets may be best suited to their needs. The ratings serve a similar purpose, but in a more compact way, allowing users to quickly see which assets are considered the best by other users. The reviews thus serve as a form of peer review. They are useful to asset providers who can use the information in them as a starting point for improving their products. They can also help users with any problems that are described in the reviews. Since the system administrators and managers also monitor the content of the system, these reviews can be helpful when performing certain system management tasks.

Monitoring Feedback

The system shall provide administrators and content managers with the ability to monitor and review all feedback about the system and assets contained in it. Feedback may include bug reports, performance, usability, and enhancement suggestions for the system and assets as well as installation and usage experiences for assets. System feedback is

helpful for maintaining the system while asset feedback is primarily of use to consumers and providers. System feedback will be reviewed and acted upon as the system managers deem necessary. Asset reviews may be monitored to make sure they are relevant to the product.

Feedback and reviews are monitored to make sure that what is available to the public is relevant and useful, not, for example, to exclude bad reviews. The system administrators and managers will review all system feedback provided to decide if it is worth acting upon, to fix problems or to improve the system. Asset reviews may be monitored to make sure they are relevant. Asset feedback will generally be reviewed by users interested in the assets. Users may have the option of challenging reviews that they feel are inappropriate in some manner and in this case, a system content manager will review the challenge and the feedback it applies to, and decide whether or not to remove the feedback.

Workflow Management

Some of the other requirements include some degree of workflow management such as approving user registrations or monitoring system feedback to make improvements as required. Other specific management features that the system shall provide include approving asset submissions and cleaning up assets. When providers submit assets to the system, a content manager will be notified that it needs to be reviewed. Upon review, useful assets with sufficient documentation and descriptions will be approved, added to the system, and made available to the public. Ones that do not meet such requirements will be rejected. In either case, the provider will be notified of the decision. Administrators can clean up the system periodically by deprecating assets which have received poor reviews, had very few downloads, or have been outdated by newer alternatives.

Such workflow management tasks are important for keeping the system running efficiently. Assets need to be approved before they are made public to make sure that they are useful products which fall into the domain covered by the system (e.g., are Earth science related in the case of a system focused on providing Earth science products). Cleaning up assets periodically helps the system evolve to meet the needs of the community by making sure that what is publicly available is useful and relevant to the community. Reasons an asset may not be used by the community include poor reviews from others or being outdated by newer alternatives. Signs of disuse include poor reviews and low ratings as well as a small number of downloads. When administrators find that assets are not being used, it is necessary to clean up the system and remove them from public access to make sure that users can quickly find the best assets for their needs.

Capture Asset Needs

While a system may provide many useful assets, it is unlikely that it will provide everything that every user is looking for. Therefore, the system shall allow administrators to determine when there is a need for an asset that the system does not contain. An administrator can review system activity and examine search results to find out what searches do not result in any matching assets. These 'no-hit' results can be analyzed to determine if there is a demand for additional assets within the community. Similarly, users shall be able to submit requests for reusable assets that currently are not in the system.

It is important to know when the system is not meeting the needs of the community so that steps may be taken to resolve the problem. If analysis of the search results show that there is a need for a particular type of asset, it may be possible for the system managers

to find it elsewhere and include it in their system. They may also announce the need on the system and encourage users to help fill this need, either by sharing an existing asset that is not yet in the system or by producing a new asset. In this way, the managers of the system can help make sure that the system continues to meet the needs of the community by providing new assets as the need for them arises.

Catalog or Repository

The basic difference between a catalog and a repository is that a repository stores the assets themselves while a catalog stores links to the assets. However, systems can combine factors of the two. For example, a system that is primarily a catalog may store assets designed and built by the organization hosting the catalog, but only those in-house items. A system that is primarily a repository should be able to display its contents to the user in the same way a catalog would. Both should provide the user with some information about each asset as well. There are advantages and disadvantages to each. A repository is a more centralized storage location, providing users with access to all available assets by keeping them on one system, but providers will have to continually submit news and updates to the repository. A catalog allows providers to do most of their own news and updates, but every asset a user wants to use is located on a different system with different methods of locating and downloading assets. Ideally, a system should be able to operate in either mode, storing either the assets themselves or links to the assets. In the case of a system designed to facilitate reuse, it should have additional features to promote the reuse of assets in the system. For example, the system might keep a record of each user's downloads and use this as a way of offering information about related assets. When a user views a particular asset, the system could note that other users who downloaded that asset also downloaded specific other assets. Providing some information about related assets that the user may find helpful encourages the reuse of more assets than just the one(s) the user specifically looks for in the system.

Minimal Operation Support

A system requires enough human support to make sure that the system runs properly and is updated on a regular basis, but this should be as little as possible. The system should be able to handle most routine tasks itself without requiring human assistance. When a system is automated as much as possible and runs properly with little intervention, fewer people are needed to support it. This saves money as well as time and effort.

Performance

The system is expected always to function correctly and to be available and operational essentially all the time. Some down time will be required for occasional maintenance, but this should be very limited. In order to make sure that the system is accessible by all users, including ones with disabilities, the system should comply with Section 508 of the Rehabilitation Act (29 U.S.C. 794d).

Security

When users are required or allowed to register with the system, it is necessary to keep their submitted information secure so that others may not obtain it. In the case of a system providing service through a web site, a secure connection can be used on all pages requiring the submission of passwords, personal information, contact details, etc. to

encrypt this information. It is also necessary that such collected information only be used in explicitly stated ways and that the user is informed of these ways before completing the registration. Network traffic may be monitored to prevent unauthorized access or changes to the system that may cause damage. A method for dealing with the submission of malicious code must also be established. One possibility is to limit upload privileges to users from reputable organizations such as NASA, other government agencies, or educational institutions. Affiliation with such a trusted organization should prevent the introduction of malicious code to the system. Once the affiliation has been checked and confirmed, upload access would be granted. In this way, only a subset of trusted users has upload privileges. The system must also be capable of meeting NASA information technology security guidelines.

Technology

In order to be easily maintainable, the system must be able to operate on standard hardware using standard software. A variety of options are available for this: commercial off-the-shelf (COTS) packages, reusable software including open source and freeware, custom developed products, etc. An appropriate mix of these products must be selected in order to produce a cost-effective and maintainable system.

Review of Existing Systems

We researched catalog and repository systems operated by NASA, as well as those outside of NASA, with a particular focus on those within the Earth science domain. All of these sites provide examples of how different features and requirements can be implemented, information that can be helpful in developing a Reuse Enablement System. The non-NASA sites are included because they were designed for similar purposes as the NASA sites. The following lists the sites that were evaluated as part of this study:

The NASA sites:

1) Global Change Master Directory (GCMD) (http://gcmd.gsfc.nasa.gov/)

2) Goddard Space Flight Center (GSFC) Open Source Software page (http://techtransfer.gsfc.nasa.gov/software.html)

3) Ames Research Center Open Source Software page (http://opensource.arc.nasa.gov/)

4) HDF-EOS Tools and Information Center (http://hdfeos.gsfc.nasa.gov/hdfeos/softwarelist.cfm)

5) Computational Technologies (CT) Project (http://ct.gsfc.nasa.gov/software.html)

6) Earth Observing System Clearinghouse (ECHO) (http://www.echo.eos.nasa.gov/)

7) Planetary Data System (PDS) Software Download at the Jet Propulsion Laboratory (http://pds.jpl.nasa.gov/tools/software_download.cfm)

The non-NASA sites:

8) Open Channel Foundation which hosts NASA's COSMIC Collection (http://www.openchannelfoundation.org/cosmic/)

9) SourceForge (http://www.sourceforge.net/)

10) Freshmeat (http://freshmeat.net/)

11) Scientific Applications on Linux (http://sal.jyu.fi/)

12) National Technology Transfer Center (http://www.nttc.edu/)

13) National HPCC Software Exchange (http://rib.cs.utk.edu/catalog.pl?rh=226)

14) Netlib (http://www.netlib.org/)

15) Savannah (http://savannah.gnu.org/)

16) Space Telescope Science Institute (STScI) Software and Hardware Products (http://www.stsci.edu/resources/software_hardware)

17) Astronomical Software and Documentation Service at STScI (http://asds.stsci.edu/)

The NASA sites were generally in the Earth science domain, but the CT Project and the GSFC Software Catalog also cover space science and the PDS Software Download is specifically space science. The non-NASA sites are generally out of the Earth science domain, but some of them do contain some software in this field. The final two sites are in the space science domain, primarily astronomy.

Feedback from some members of the Earth science community provided information on some other sites used to locate reusable assets. These sites included:

18) Direct Readout Laboratory (http://directreadout.nasa.gov/)

19) Glenn Research Center Software Repository (https://technology.grc.nasa.gov/software/cat_list.asp?Center=GRC)

20) ArcScripts (http://arcscripts.esri.com/)

21) Wikipedia (http://www.wikipedia.org/)

22) Usenet newsgroups (e.g., http://groups.google.com/)

23) Ruby Application Archive (http://raa.ruby-lang.org/)

24) SciRuby (http://sciruby.codeforpeople.com/)

25) Comprehensive Perl Archive Network (http://www.cpan.org/)

26) FreeGIS (http://www.freegis.org/)

These sites were examined, but not reviewed in detail. A brief description of each follows the detailed review of the other sites listed above.

This section describes the result of our study on how well these sites meet the requirements defined in the previous section. Where appropriate, the system's ability to meet our requirements is rated on a scale of 0 to 3 stars with 0 being no support, 1 being some support, 2 being good support, and 3 being full support. Due to the nature of this study, we are able to review the systems only from a consumer viewpoint. We are unable to determine how certain administrative functions are handled and therefore the requirements regarding workflow management, capturing asset needs, and performance have been omitted from our detailed review. The domain of the system and the type of assets provided, while not functional requirements, are important factors in determining whether or not the system meets the needs of the Earth science community, so these points have been included in the detailed review. The number of items available at each site was current at the time the sites were reviewed. Updates since then may have increased the total number of items available at some sites.

Global Change Master Directory (GCMD)

The GCMD is owned by NASA and is run by the Global Change Data Center within the Earth Sciences Directorate at the Goddard Space Flight Center. Its goal is "to enable users to locate and obtain access to Earth science data sets and services relevant to the global change and Earth science research."

Site registration is not required or offered and all visitors can view the site and the contents of its catalog. Users are able to contribute assets by submitting the relevant information through a web-based form interface or stand-alone tool. System feedback is accepted through a web form. Automatic notifications are based on keywords for location and/or Earth science areas and available to all users who will receive e-mail updates once a month. Assets can be located by hierarchy, search, and a combination of both. Lists of all assets within a hierarchy category are also available. Users cannot register asset usage or provide asset reviews. Feedback monitoring is limited to system feedback. The current catalog consists of approximately 16000 data sets and 1240 data services.

The GCMD has very good coverage of the Earth science domain, especially in the data set hierarchies; the data services are categorized differently. Lists within the hierarchy categories are ordered by date, making the newest assets the most prominent; free-text searches are ranked by word occurrence. It was not always easy to find known assets because of the specific sub-categories they were classified under. Using the advanced search from the link on the data sets and data services pages allows the user to select certain field(s) on which the search will operate. This can be better than using the simple search box, but tips for using the free-text search are provided. Contributed assets need to be in particular formats, but a web-based interface simplifies this procedure so that users do not necessarily need to know the details of the formats involved. The ease of downloading depends on the external site hosting the asset, but uploading new asset metadata appeared to be simple thanks to the web-based form and interface. Support is available by web form and individual asset contacts. The site functions primarily as a catalog of metadata, but does have the ability to host some very small data sets.

Requirement/Feature	Available at GCMD?

Requirement/Feature	Available at GCMD?
Domain	Earth science
Type of Assets	Metadata to access data setsand data services
Register User	☆☆☆
Contribute/Update Assets	★★☆
Provide System Feedback	★★☆
Automatic Notifications	★★★
Discovering Assets	Hierarchy, Search
Register Asset Usage	☆☆☆☆
Provide Asset Review	☆☆☆☆
Monitoring Feedback	★☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Catalog
Operation Support	Ten member staff includes four developers
Technology	RSYNC, Zope, CVS, Linux, Java, JavaServer Pages, XML, Apache, Oracle/PostgreSQL, Struts, Lucene, XSLT, Dreamweaver

The GCMD primarily provides data sets and services and not many reusable software components, but does have good coverage of the Earth science domain and offers links to a very large number of products. However, it lacks the ability to register users, register asset usage, or provide asset reviews. Data sets and services are provided, but the types of assets are variable; data sets include papers and data services include links to web-based interfaces for other systems. Software appears to be finished products in most cases, but since the GCMD functions as a catalog, visiting the external sites hosting the software is necessary to determine exactly what is available. It may be difficult for software developers to reuse such packages when developing new products.

Goddard Space Flight Center (GSFC) Open Source Software

The GSFC Open Source Software catalog is owned by NASA and run by the Office of Technology Transfer at the Goddard Space Flight Center. Its goal is to distribute software produced by NASA and released under the NASA Open Source Agreement to "accelerate NASA software development and improve the quality via community peer review,

maximize the awareness and impact of NASA research, and increase dissemination of NASA software in support of NASA's education mission."

All site visitors are allowed to view the site and its contents; site registration is not available. As the site is designed for the release of open source software developed at GSFC, general users cannot submit assets. System feedback may be provided by e-mail to the web master or the office in charge of the site. Users can register asset usage and this allows them to receive automatic notifications of updates to the products they register. Assets can be located by a complete list. Asset reviews may be provided as part of the asset registration form, but these comments do not appear to be published on the site. Feedback monitoring is assumed to occur for both system feedback and the comments users can submit during asset registration. The catalog/repository currently contains 11 entries.

This catalog covers the science domain in general, but has some programs specific to Earth and space science. However, it is quite small – there are seven open source software projects (one of which is ECHO, reviewed here), one freeware project, and three licensable software projects. The open source projects are normally distributed as source code. Downloading is very easy for software provided on the site, but some items must be obtained from external sites. Support is available by e-mail and phone.

Requirement/Feature	*Available at GSFC Catalog?*
Domain	Earth and space science
Type of Assets	Open source packages
Register User	☆☆☆
Contribute/Update Assets	★☆☆
Provide System Feedback	★★☆
Automatic Notifications	★☆☆
Discovering Assets	List
Register Asset Usage	★★★
Provide Asset Review	★☆☆
Monitoring Feedback	★★☆
Secure Log In / Registration	N/A
Catalog or Repository	Both
Operation Support	Uncertain, presumed small
Technology	PHP

This site only distributes NASA-produced open source software, placing a strong restriction on what is available. It meets some of the requirements, but is lacking the ability to register users and to allow users to contribute assets. Also, the register asset usage page is not secure. Having only a list works for the small size of this catalog, but hierarchies and searches are generally better options. Users can submit reviews and comments when they register an asset, but these reviews are not posted on the site, so new users are unable to benefit from the experience of past users. Distributing open source software is a positive point, but these are typically finished products that could be difficult for software developers to reuse when creating new assets.

NASA Ames Research Center Open Source Software

The Ames Open Source Software catalog is owned by NASA and run by the Ames Research Center. Its goal is to distribute software produced by NASA and released under the NASA Open Source Agreement to "accelerate NASA software development and improve the quality via community peer review, maximize the awareness and impact of NASA research, and increase dissemination of NASA software in support of NASA's education mission."

All site visitors are allowed to view the site and its contents; site registration is not available. As the site is designed for the release of open source software developed at NASA, general users cannot submit assets. System feedback may be provided through a web-based form. Users can register asset usage and this allows them to receive automatic notifications of updates to the products they register. Assets can be located by a complete list. There is no apparent way to provide asset review. Feedback monitoring is limited to system feedback. The catalog/repository currently contains 9 entries.

This catalog covers the science domain in general, but has a few assets specific to Earth science. However, it is quite small – there are nine open source software projects, one of which is ECHO (reviewed here) which are normally distributed as source code. Downloading is very easy and the optional asset registration form is provided on the download page. This page is not secure though. Support should be available by e-mail or web-based form, but the feedback link on the page is broken; the general "contact Ames" link on the page it leads to must serve as an alternative.

Requirement/Feature	Available at Ames OSS?
Domain	General science
Type of Assets	Open source packages
Register User	☆☆☆
Contribute/Update Assets	★☆☆
Provide System Feedback	★★☆
Automatic Notifications	★☆☆
Discovering Assets	List
Register Asset Usage	★★★

Requirement/Feature	Available at Ames OSS?
Provide Asset Review	☆☆☆
Monitoring Feedback	★☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Both
Operation Support	Uncertain, presumed small
Technology	JavaServer Pages

This site only distributes NASA-produced open source software, placing a strong restriction on what is available. It meets some of the requirements, but is lacking the ability to register users, to allow users to contribute assets, and to provide asset reviews. Also, the register asset usage page is not secure. Having only a list works for the small size of this catalog, but hierarchies and searches are generally better options. Distributing open source software is a positive point, but these are typically finished products that could be difficult for software developers to reuse when creating new assets.

HDF-EOS Tools and Information Center

The HDF-EOS Tools and Information Center is owned by NASA and run by the Goddard Space Flight Center. Its goal is to provide software that supports HDF or HDF-EOS.

All visitors are allowed to view the site and its contents and site registration is not required. However, in order to download software, a very basic form must be completed with name, e-mail address, and country, but this page containing this form is not secure. User-submitted software is posted, so submissions are accepted, but this must be handled via e-mail as there is no submission form on the web site. System feedback may be provided in the site's web forum or by e-mail to the web master. Automatic notifications are not available. Assets may be located by complete list, hierarchy, and filtering of the list. Users cannot register asset usage; the registration used here only tracks downloads, not active usage as in our requirements. Asset reviews may be provided by ratings (from 1-5 on installation, features, stability, and usage) and comments on the individual assets and through the site's web forum; anonymous reviews are allowed. Feedback is monitored in the sense that submitted reviews do not appear immediately, but are sent to the web site administrator and posted at a later date. There are currently 53 tools available in the site's repository and 17 of them have been rated by users.

The domain here is Earth science related, but very specific in that it only contains tools that support HDF and HDF-EOS, the Hierarchical Data Format used by all NASA Earth Observing System data products. These are typically finished products, sometimes with source code, and some of them are commercial products. While there is no regular search option, filtering the list based on hierarchy category, platform, and/or data application can serve as a substitute. Downloading was simple and support is available by e-mail, forum, and individual asset contacts.

Requirement/Feature	Available at HDF-EOS Tools?
Domain	Earth science, HDF/HDF-EOS
Type of Assets	finished products
Register User	★☆☆
Contribute/Update Assets	★★☆
Provide System Feedback	★★★
Automatic Notifications	☆☆☆
Discovering Assets	List, Filter
Register Asset Usage	★☆☆
Provide Asset Review	★★★
Monitoring Feedback	★★☆
Secure Log In / Registration	NO
Catalog or Repository	Repository
Operation Support	None, site is inactive
Technology	Cold Fusion

The HDF-EOS Tools and Information Center contains only HDF and HDF-EOS related software, providing a very narrowly focused site. It does have many of the features defined in the requirements here, but is missing the ability to register users, provide automatic notifications, and register asset usage (downloads are registered instead). Contributing assets cannot be done automatically and the degree of feedback monitoring for asset reviews is unclear (it is assumed that system feedback is monitored). Most of these tools are finished products which may be difficult for software developers to use in creating new products. While this does cover the Earth science domain, it is so specific to HDF and HDF-EOS tools that it cannot meet the needs of the community as a whole.

In addition, the site has closed down completely since it was reviewed, so it is no longer available.

Computational Technologies (CT) Project

The CT Project is owned by NASA and run by the Goddard Space Flight Center. Its aim is to make "teraFLOPS (trillion floating-point operations per second) computing systems an integral part of large-scale computing activities in the Earth and space sciences."

All visitors are permitted to view the site and its contents and there is no site registration available. Asset contributions are not accepted as this site maintains information about

the Grand Challenge projects and only their products are listed. System feedback may be given by e-mail to the web master. Automatic notifications are not available. Assets may be located by a sort of combination hierarchy/list – after selecting which Grand Challenge to view, all projects within in it are listed in categories. A search box is provided, but it searches within the entire domain of the site, not specifically within the software section. There is no feature to provide asset registration or review and feedback monitoring is limited to system feedback. Approximately 40 projects are listed in the catalog.

This site covers both Earth and space science. The hierarchy is rather unstructured – the top level is by Grand Challenge Investigator group and within them, each individual project gets its own category. As noted above, the search box provided searches the whole site, not just the software section, so users must explore the site to find what they want, if it is even available here. Since the site functions as a catalog rather than a repository, users must follow links to external sites to obtain assets. Some direct links to software are provided on the site, but at least some of these are broken, leaving the user to search the external sites. It was not apparent that software is readily available from all projects, but for active links with software, source code is usually distributed. Some of the software links provided on the CT Project pages were broken and it was not always easy to locate software on the external pages for each project. Support is available by e-mail and individual asset contacts.

Requirement/Feature	Available at CT Project?
Domain	Earth and space science
Type of Assets	finished products with source code
Register User	☆☆☆
Contribute/Update Assets	☆☆☆
Provide System Feedback	★★☆
Automatic Notifications	☆☆☆
Discovering Assets	Hierarchy
Register Asset Usage	☆☆☆
Provide Asset Review	☆☆☆
Monitoring Feedback	★☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Catalog
Operation Support	Uncertain, presumed small
Technology	HTML?

The catalog at the CT Project site meets few of the requirements defined here. Part of this is probably due to its nature as a catalog – by always linking to external sites, it tends to put more emphasis on the individual sites providing the features. Those sites may meet the requirements, but it would be done inconsistently. It does cover the Earth and space science domain, but with a limited number of specific projects normally producing finished products, so it appears to be of limited use in providing reusable assets that software developers could use in their own projects.

Earth Observing System Clearinghouse (ECHO)

The ECHO system is owned by NASA and run by NASA's Earth Science Data and Information System. Its goal is to support "efficient discovery and access to collection- and inventory-specific Earth science data and services."

All visitors can view the site, but accessing the metadata stored by this project requires the use of a client. There is no site registration, but an application must be submitted and approved in order to become a partner (data, client, or service). Partners are allowed to contribute assets and are also expected to have certain levels of computer experience stated on the pages for each type of partner. System feedback may be provided via e-mail and an API that supports subscription for collection level information is available. Assets may be located by search using a client. There is currently no way to register asset usage or provide asset reviews and monitoring feedback is limited to system feedback. The clearinghouse/catalog currently contains 1942 collections (approximately 50 million granules).

The ECHO project is in the Earth science domain, but is designed to provide access to metadata, allowing the exchange of data and information between members of the community. It serves a somewhat different purpose than the rest of the sites we examined and does not appear to contain software. Only one ECHO client is listed as operational, but six others are under development. The operational client provides keyword, spatial and temporal search choices, but there was no obvious way to download an item in the search results. Support is available by e-mail, phone, and mailing lists. The Warehouse Inventory Search Tool (WIST) is a general-purpose client that will replace the EOS Data Gateway for accessing EOS and related data. Currently, a beta version is available that provides a basic, end-to-end capability for searching and data access.

Requirement/Feature	Available at ECHO?
Domain	Earth science
Type of Assets	Metadata for data and services
Register User	★☆☆
Contribute/Update Assets	★★☆
Provide System Feedback	★★☆
Automatic Notifications	☆☆☆

Requirement/Feature	Available at ECHO?
Discovering Assets	Search
Register Asset Usage	☆☆☆
Provide Asset Review	☆☆☆
Monitoring Feedback	★☆☆
Secure Log In / Registration	YES
Catalog or Repository	Catalog
Operation Support	Available
Technology	XML (WSDL), API, SOAP, UDDI

ECHO is primarily a data/service broker, not a software provider and the site is missing a number of the defined requirements. Assets may be added by partners, but registering to become a partner is apparently handled off line. The site requires partners to have a certain level of experience with the technology used on the site in order to be able to add assets. ECHO is essentially middleware; it is a clearinghouse for metadata and is not designed to function as a reuse repository.

Planetary Data System (PDS) Software Download

The PDS Software Download repository is owned by NASA and run by the Jet Propulsion Laboratory. Its goal is to provide software tools "to facilitate data capture, archiving, analysis and manipulation."

All users can visit the site and view its contents. Site registration is not required, but a basic web-based form must be completed in order to download assets. There does not appear to be any way for general users to contribute assets. System feedback may be provided by e-mail to the curator or a general "comments and questions" address. Automatic notifications are not provided. Assets may be located through a complete list. Users cannot register asset usage; the registration used here only tracks downloads, not active usage as in our requirements. The registration page is not secure. There is no apparent way to provide asset reviews and feedback monitoring is limited to system feedback. There are currently 39 entries listed in this repository.

The site is easy to navigate since there is only a list of assets. Some assets require contacting PDS for assistance and some are stored on external sites. For the ones directly available on the site, downloading seems to be easy and only requires completing a basic license agreement form. A number of the listed assets are available as part of a larger package and apparently only available that way. Despite being listed separately, they can not be downloaded separately. The assets are generally tools that are distributed as binaries and source code. The domain of the system is planetary astronomy. Support is available by e-mail to the curator or a general help address, but individual assets hosted on external sites may have other means of support.

Requirement/Feature	Available at PDS?
Domain	Planetary astronomy
Type of Assets	Mainly tools, binaries and source
Register User	☆☆☆
Contribute/Update Assets	☆☆☆
Provide System Feedback	★★☆
Automatic Notifications	☆☆☆
Discovering Assets	List
Register Asset Usage	☆☆☆
Provide Asset Review	☆☆☆
Monitoring Feedback	★☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Both
Operation Support	Uncertain, presumed small
Technology	Cold Fusion

The PDS Software Download site is for planetary astronomy, not Earth science. This system meets few of the requirements, providing little beyond the ability to provide and monitor system feedback. Many of the assets are tools rather than complete finished products and these smaller sized components are more easily reused. Some tools are only available as part of a larger package and others require contacting PDS for more information. Not all assets are hosted on the site, some reside on external sites. Also, the domain is planetary astronomy, a particular area within space science.

COSMIC Collection / Open Channel Foundation (OCF)

The COSMIC Collection is owned by NASA and is run by the Open Channel Foundation. Its goal is to provide software "created by NASA in a wide range of disciplines including engineering, chemistry, aerodynamics, and other areas."

The COSMIC Collection is hosted at the Open Channel Foundation (OCF) and allows all visitors to view the site and its contents, but users must register an account before they can download software; the log in is secure. This particular collection is apparently static and will not accept new contributions, but the OCF in general does accept submissions of new assets. System feedback is handled through a web form for submitting bugs. Automatic notifications are available to registered users who may choose to monitor software titles. Assets can be located via complete list, hierarchies, and search, but there

is no ability to register the usage of an asset. There is no clear facility for providing asset reviews. If asset moderators set up forums, reviews could be posted there, but this is not a central location designed specifically for reviews. The moderators for each asset are responsible for monitoring the feedback received for that asset. Only 192 of the more than 500 entries in the COSMIC Collection are currently available in the site's repository, but users can request software that does not appear and the OCF will attempt to get access to the program for them.

The COSMIC Collection covers a wide range of disciplines, so the amount of Earth science material available is limited; the same is true for the OCF in general. In order to obtain any software, a request form must be completed on line and a license form must be printed out, signed, and then returned to OCF by FAX or e-mail, sending a scan of the document. When required, payment is handled by the web-based request form. Obtaining software is therefore not instantaneous; the delay depends on how OCF handles requests. The assets are generally finished products, often with source code available, and the cost can be very expensive, particularly to obtain the source code. Support is available by e-mail / form, forum, and individual asset pages (in some cases).

Requirement/Feature	*Available at OCF / COSMIC?*
Domain	Wide range of disciplines
Type of Assets	Finished products with source code
Register User	★★★
Contribute/Update Assets	OCF: ★★☆ COSMIC: ☆☆☆
Provide System Feedback	★☆☆
Automatic Notifications	★★☆
Discovering Assets	List, Hierarchy, Search
Register Asset Usage	☆☆☆
Provide Asset Review	☆☆☆
Monitoring Feedback	★★☆
Secure Log In / Registration	YES
Catalog or Repository	Repository
Operation Support	Five member Board of Advisors
Technology	Apache, PHP, MySQL

The COSMIC Collection at OCF is a general science/engineering site and not Earth science specific. It meets a number of the requirements defined here, but users cannot

register asset usage or submit assets to the COSMIC Collection, submissions to the OCF in general are accepted though, and there is no dedicated asset review feature. Also, a few of the points are applied inconsistently. Providing asset reviews is possible, but the availability of this feature depends on how the monitors of each individual asset have set up the page for that asset. If they provided a forum, then reviews can be provided there, but it will contain other information, discussion of problems, etc. as well. Similarly, the extent to which feedback is monitored depends on the moderators of each individual asset. The OCF and COSMIC cover a wide range of disciplines meaning there is a limited amount of Earth science material available. Since mostly finished products are provided, even with the source code, these may be difficult for other software developers to reuse when creating new products. There are also complications involved in having a separate organization run the repository, primarily because all changes must go through and be performed by them.

SourceForge

SourceForge is owned and run by the Open Source Technology Group, part of the VA Software Corporation. Its goal is to provide free services to software developers, operating as a "development and download repository of open source code and applications."

All visitors can view the site and its contents and download assets. Site registration is offered and is typically used to provide additional services, but it is required before contributing a new asset or joining the development of an existing one. The log in is secure. System feedback may be provided through the site's tracker system via web-based form. Automatic notifications are available through a monitoring system for registered users, but appear to be for subsets of information on each project; monitoring projects as a whole requires paying for a site subscription. Assets can be located through multiple hierarchies and search. There does not appear to be a way to register asset usage as defined here. There is no specific feature designed for handling asset reviews. This feedback could be posted in forums for the assets, but could get lost among the general discussion taking place there. There is no explicit indication of the level of feedback monitoring, but it is assumed that all system feedback is monitored and the individual asset administrators will monitor any asset specific feedback that is received. There are approximately 102 000 registered projects at this repository.

This is a general software repository, covering all domains, so there is a limited amount of Earth science material available. All of the projects offered are open source software, typically finished products, and generally distributed in binary form and source code. A variety of top-level hierarchy categories are available, making it easier to find desired assets, and some degree of filtering can be applied to lists within the hierarchies as well. Downloading is simple and a sidebar lists the ten most downloaded and ten most active projects, making it easy to find the most popular assets. Support is available by web form and project contacts or the other features available on individual project pages (e.g., forum, mailing list, tracker).

Requirement/Feature	Available at SourceForge?
Domain	General software repository
Type of Assets	Open source packages

Requirement/Feature	Available at SourceForge?
Register User	★★★
Contribute/Update Assets	★★★
Provide System Feedback	★☆☆
Automatic Notifications	★★☆
Discovering Assets	Hierarchy, Search
Register Asset Usage	☆☆☆
Provide Asset Review	★☆☆
Monitoring Feedback	★★☆
Secure Log In / Registration	YES
Catalog or Repository	Repository
Operation Support	Eleven full-time staff members
Technology	PHP

SourceForge is a general software repository with no focus on Earth science. Most of the requirements are met, but it is lacking the ability to register asset usage and to provide asset reviews, at least in one clear location which makes the information difficult to find, if present. While automatic notifications are available, they seem to be somewhat complicated, tracking certain subsets of information within a project. Monitoring entire projects requires a paid subscription and apparently tracks all such subsets automatically. It is unclear if there is a way to obtain, for example, automatic notification that a new version of the software has been uploaded. As a general software repository, the amount of software available in any one specific field is typically a small fraction of the total – in the Scientific/Engineering category, Earth Science has 29 projects, Ecosystem Sciences has 13 projects and GIS has 228 projects. Without focusing on a specific field, it will be difficult to meet the needs of that field. The distribution of open source software is good, but as usual, the assets being offered are typically final packages designed for the end user and not the smaller components that would be more beneficial to software developers creating new assets.

Freshmeat

Freshmeat is owned and run by the Open Source Technology Group, part of the VA Software Corporation. Its goal is to provide "Unix and cross-platform software, themes and related 'eye-candy', and Palm OS software", preferably released under open source licenses.

The site can be viewed by all visitors and they are able to download assets from the repository. Site registration is provided and mainly used to provide users with additional

features, but it is required before uploading new projects. The log in is not secure. System feedback may be provided through a web-based form. Automatic notification is available for registered users. Assets may be located through multiple hierarchies and search. There is no apparent way to register asset usage. Asset feedback may be provided by comments and/or ratings; both features require that the user be registered and logged in. System feedback is assumed to be monitored, but it is unclear how much asset feedback is monitored; that is probably dependent on the individual project administrators. There are approximately 37 700 projects in this catalog.

As a general software catalog, there is a limited amount of Earth science material available here – there are only 76 projects in the Scientific/Engineering, Geographical category. The hierarchy indicates that about 28 600 projects are released under an Open Source Initiative approved license and the stats page shows that about 73% of project branches are released under either the GNU General Public License (GPL) or the GNU Lesser General Public License (LGPL). However, not all software is free and the focus of the site is on software for Unix-like operating systems. A variety of top-level hierarchy categories make it easier to find assets. Even though the assets are on external pages, downloading them is simple since direct download links are provided. A stats page provides lists of the top twenty projects by popularity, vitality, and rating making the most popular projects easy to find. Support is available through a web-based form for system issues and through individual project contacts for asset issues; individual assets may provide other means of support, such as a mailing list, as well.

Requirement/Feature	Available at Freshmeat?
Domain	General software catalog
Type of Assets	Mainly open source packages
Register User	★★☆
Contribute/Update Assets	★★☆
Provide System Feedback	★★☆
Automatic Notifications	★★☆
Discovering Assets	Hierarchy, Search
Register Asset Usage	☆☆☆☆
Provide Asset Review	★★★
Monitoring Feedback	★★☆
Secure Log In / Registration	NO
Catalog or Repository	Repository
Operation Support	One person on development Six people on projects

Requirement/Feature	Available at Freshmeat?
Technology	XML-RPC?

Freshmeat is a general software repository with no focus on Earth science. The system satisfies most of the requirements, but it lacks the ability to register asset usage and does not provide a secure log in. There is no uniform monitoring of asset feedback, this apparently being left to the individual project administrators to handle. As a general software catalog, the amount of software available in any specific field is limited. There is a Geographical category here, as noted above, but no Earth Science category. Most of the software is open source, but not all projects are free. The focus on Unix and cross-platform software could limit the usefulness of this resource for some potential users. The packages provided are typically aimed at the end user which could make it difficult for software developers to reuse assets in their own projects.

Scientific Applications on Linux (SAL)

SAL is owned and run by Kachina Technologies, Inc. Its goal is to provide "a collection of information and links to software that will be of interest to scientists and engineers."

All visitors to the site may view the contents of the catalog and follow links to the external sites hosting the assets. Site registration is not provided. Asset contributions are accepted, but there is no automated way to submit them, and the site is apparently inactive, so this may no longer be true. System feedback may be provided through a web-based form. Automatic notifications do not appear to be available as defined here, although a general discussion mailing list which is also used to announce new additions is listed. Assets may be located via hierarchies or search. There is no way to register asset usage. Comments are accepted and may be used for asset review, but there is no automated way to submit them on the web site. System feedback is assumed to be monitored and asset feedback should be monitored, at least to the extent that it must be posted to the site by an administrator. There are 3070 entries listed in this catalog.

The major issue with this site is that it appears to be inactive. The last copyright on the page is 2001, links to mirror sites are often broken, and links to Kachina Technologies, Inc., the company that managed the site together with the original creator, redirect to a scholarship related web site. While support may have originally been available, it appears that support cannot be obtained now. The hierarchy structure is offered through a set of pull-down menus, but is still simple to navigate through. The search option appears to be broken though. The domain is science in general, but it still has a limited amount of Earth science material available – 27 assets are listed under Other Scientific Fields, Misc., Sciences of Earth/Environmental/Ocean. In general, assets are a mix of commercial, shareware, GPL, and other licenses. Ease of downloading depends on the external site hosting the software.

Requirement/Feature	Available at SAL?
Domain	Scientific software catalog

Requirement/Feature	Available at SAL?
Type of Assets	Mainly tools and packages with source code
Register User	☆☆☆
Contribute/Update Assets	* ☆☆☆
Provide System Feedback	* ☆☆☆
Automatic Notifications	* ☆☆☆
Discovering Assets	Hierarchy, Search (broken)
Register Asset Usage	☆☆☆
Provide Asset Review	* ☆☆☆
Monitoring Feedback	* ☆☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Catalog
Operation Support	Unknown, probably inactive
Technology	HTML, Java?

This site has become inactive, so many of the requirements have been marked with asterisks (*) to indicate that these points may have been addressed when the site was actively maintained. The catalog covers science in general, but still contains a limited number of Earth science related assets. The assets offered are typically final tools and packages aimed at the end user, so even with the source code available, they may be difficult for software developers to reuse when creating new assets

National Technology Transfer Center (NTTC)

The NTTC is owned and run by the National Technology Transfer Center based at the Wheeling Jesuit University. Its goal is to "promote technologies and technology needs from our federal clients", the majority of which is produced by NASA.

All site visitors may view the site and its contents. Site registration is not offered and direct downloads are not available. In order to obtain any software, the user must complete a web form to receive more information and these submission pages are not secure. There does not appear to be any way for general users to contribute assets to the site. System feedback may be provided by e-mail to the web master. Automatic notifications of new assets being added are available and are based on technology categories. Assets may be located by complete list, hierarchies, and search. There is no apparent way to register asset usage or to provide asset review. Feedback monitoring is limited to system feedback. There are 184 technologies listed in the catalog.

The hierarchy and search functions can be used together, as the hierarchy lists are available from a pull-down menu within the search feature. Together they allow searching within a specific category. The site appears to be a go-between for users and providers. Apparently the listed technologies area available, but they are not provided on the site. Users must submit forms declaring their interest in a product before anything else can be done. These forms require the submission of some personal information (e.g., mailing address and phone number), but they are not on secure web pages. The majority of the technologies listed are from NASA, but there seems to be a limited amount of Earth science products. Support is available by e-mail to the web master.

Requirement/Feature	Available at NTTC?
Domain	Federal technology, but most comes from NASA
Type of Assets	Finished products?
Register User	☆☆☆☆
Contribute/Update Assets	★☆☆
Provide System Feedback	★★☆
Automatic Notifications	★★☆
Discovering Assets	List, Hierarchy, Search
Register Asset Usage	☆☆☆☆
Provide Asset Review	☆☆☆☆
Monitoring Feedback	★☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Both?
Operation Support	Uncertain
Technology	Active Server Pages

This system meets few of the requirements defined here because it is somewhat different from a typical catalog or repository. Technologies are listed, as in a catalog, and are presumably available, as in a repository, but users cannot obtain them directly. They must complete forms to receive more information first. These forms require some personal information, but are not handled on secure web pages. This process delays the user from obtaining the desired product. The offered technologies are assumed to be final packages designed for end users and these can be difficult for software developers to reuse when creating new assets.

National HPCC Software Exchange Meta-Repository (NHSE)

The NHSE is owned by the High Performance Computing and Communication agencies and developed by the Center for Research on Parallel Computation. It is "a distributed collection of software, documents, data, and information of interest to the high performance and parallel computing community."

All users of the site can view the contents of the catalog and follow links to the external sites hosting the assets. There is no obvious way to contribute assets and since the site is no longer actively maintained, contributing may no longer be possible. System feedback can be provided by e-mail to the repository contact listed on the "About NHSE" page of the site. Automatic notifications are not available. Assets may be located by hierarchy and search. There is no way to register asset usage and no facility for providing asset reviews. Monitoring feedback is limited to system feedback. This catalog contains 265 entries.

A major issue with this site is that lack of funding has made continuing operation of the site impractical for NHSE. This system is aimed at the high performance and parallel computing community, so there is a limited amount of Earth science assets available. The site is easy to navigate, but not all top-level hierarchy categories have sub-categories and not all sub-categories have contents. Not all items in the category lists have short descriptions provided, so sometimes going to the asset page is required to find out what purpose it serves. Ease of downloading depends on the external site hosting the asset. Support may be available by e-mail; external sites hosting individual assets may also have their own support.

Requirement/Feature	Available at NHSE?
Domain	High performance and parallel computing
Type of Assets	Tools and finished products
Register User	☆☆☆☆
Contribute/Update Assets	* ☆☆☆☆
Provide System Feedback	* ☆☆☆☆
Automatic Notifications	☆☆☆☆
Discovering Assets	Hierarchy, Search
Register Asset Usage	☆☆☆☆
Provide Asset Review	* ☆☆☆☆
Monitoring Feedback	* ☆☆☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Catalog

Requirement/Feature	Available at NHSE?
Operation Support	None, site is inactive
Technology	Repository In a Box

Support for this site has been withdrawn due to lack of funding, so a few of the requirements above have been marked with asterisks (*) to indicate that they may have been addressed when the site was actively maintained. In its current state, it does not meet most of the requirements. It does seem to provide more smaller-sized components than most other sites which are beneficial for reuse by software developers, but finished products still appear to be the primary type of asset provided. Also, there is a limited amount of Earth science related material since the domain is high performance and parallel computing.

Netlib

Netlib is a repository maintained by AT&T Bell Laboratories, the University of Tennessee, and Oak Ridge National Laboratory with assistance from colleagues world-wide. It contains "freely available software, documents, and databases of interest to the numerical, scientific computing, and other communities."

All users can view the site and its contents and download assets from it. Site registration is not provided. Contributions are accepted, but there is no automated way to submit assets on the web site; users must e-mail the site maintainers first and proceed from there. System feedback may be offered by e-mail to the maintainers. There is no automatic notification feature on this site. Assets may be located by hierarchy or search. There is no apparent way to register asset usage or to provide asset reviews. Monitoring feedback is limited to system feedback. The site does not provide a count of the number of assets available, but a few trial searches indicate that there are hundreds of items available in this repository.

The site is easy to navigate, but the hierarchy categories listed on the browse page have no descriptions, so the contents of each category can be unclear. A link to a similar top-level hierarchy listing with descriptions of the categories is provided. The site contains papers and databases in addition to software and much of the software is provided as computer code. These are frequently smaller components like functions that could be incorporated into larger software products. However, the domain is mathematics and there are essentially no Earth science specific assets. Support is available by e-mail and a forum, though posting to the forum is currently disabled.

Requirement/Feature	Available at Netlib?
Domain	Mathematics
Type of Assets	Mainly source code for functions and similar sized components
Register User	☆☆☆☆

Requirement/Feature	Available at Netlib?
Contribute/Update Assets	★★☆
Provide System Feedback	★★☆
Automatic Notifications	☆☆☆
Discovering Assets	Hierarchy, Search
Register Asset Usage	☆☆☆
Provide Asset Review	☆☆☆
Monitoring Feedback	★☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Repository
Operation Support	Total 17 editors, administrators, and developers
Technology	HTML?

Netlib contains mathematics software which may or may not be applicable to Earth science. This system meets some of the requirements, but lacks the ability to register users, provide automatic notifications, register asset usage, or provide asset reviews. It offers a number of smaller sized software components, such as functions, and this is very beneficial to software developers who can include the provided source code in their projects. Finished products are also provided and some of the components in these packages may be reusable in other projects as well. In some cases, the components can be downloaded independently without obtaining the full package. These are good points, but the domain of the site is not Earth science.

Savannah

Savannah is owned and run by the Free Software Foundation, Inc. Its goal is to provide "a central point for development, distribution and maintenance of GNU Software."

All users may view the site's contents and download assets. Site registration is available, but only required in order to contribute assets and join in the development of existing ones. The log in page is secure. System feedback may be provided through a web-based form or e-mail. An automatic notification system does not appear to be in place for the site as a whole, but individual projects could send updates through mailing lists, when available. Assets can be located by complete list and search. There is no apparent way to register asset usage and there is no specific feature designed for handling asset reviews. This feedback could be posted in forums for the assets, but could get lost among the general discussion taking place there. All system feedback is assumed to be monitored while individual project administrators are responsible for monitoring feedback specific to each asset. There are approximately 2410 entries in this repository.

The lack of a real hierarchy makes it harder to locate software designed to perform a particular task; the categories that do exist are, for example, Official GNU Software and Non-GNU Software & Documentation, which are too broad to be of much use. A sidebar on the main page lists the newest projects in these categories, making recent additions easy to find; some statistics are provided in this sidebar as well. The domain is general software with some emphasis placed on official GNU software. Support is available through web-based form and individual project pages.

Requirement/Feature	Available at Savannah?
Domain	General
Type of Assets	Usually tools and packages
Register User	★★☆
Contribute/Update Assets	★★☆
Provide System Feedback	★★☆
Automatic Notifications	★★☆
Discovering Assets	List, Search
Register Asset Usage	☆☆☆
Provide Asset Review	☆☆☆
Monitoring Feedback	★★☆
Secure Log In / Registration	YES
Catalog or Repository	Repository
Operation Support	15 administrators for the software hosting system
Technology	Apache, Perl, PHP, MySQL

As a general software repository, Savannah has a limited amount of Earth science material available. This system meets some of the requirements, but lacks automatic notifications and the abilities to register asset usage and to provide asset reviews. There is no real hierarchy system which can make it harder to find a specific type of asset. Available assets seem to be mainly tools and packages, both primarily aimed at the end user. This type of software can be difficult for software developers to reuse in their own projects.

Space Telescope Science Institute (STScI) Software and Hardware

The STScI Software and Hardware catalog is owned and run by the Space Science Telescope Institute. Its goal is to provide software and hardware resources for use in astronomy.

All users may view the site and its contents and download assets from it. As this site is designed to host software developed at STScI, general users cannot submit new assets to the collection. System feedback may be provided by e-mail to a general help address. There is no automatic notification feature. Assets may be located by list and hierarchy. There is no apparent way to register asset usage or provide asset reviews. Feedback monitoring is limited to system feedback. There are currently 14 entries in the repository.

The domain here is astronomy and thus in space science rather than Earth science. The full list groups the assets by hierarchy category and the hierarchy is also available in a separate sidebar. For such a small repository, this is somewhat redundant, but as the repository grows, the separate hierarchy list will become more useful than the full list of assets grouped by category. Some of the categories in the sidebar are empty. The provided assets seem to be a mix of stand-alone, plug-in, and modules, often distributed as source code. Downloading is generally easy, but some assets require contacting STScI for more information first, meaning that it will take longer to obtain them. Support is available by e-mail to a general help address.

Requirement/Feature	Available at STScI?
Domain	Astronomy
Type of Assets	Packages, usually with source
Register User	☆☆☆
Contribute/Update Assets	☆☆☆
Provide System Feedback	★★☆
Automatic Notifications	☆☆☆
Discovering Assets	List, Hierarchy
Register Asset Usage	☆☆☆
Provide Asset Review	☆☆☆
Monitoring Feedback	★☆☆
Secure Log In / Registration	N/A
Catalog or Repository	Repository
Operation Support	Uncertain, presumed small
Technology	HTML?

This system lies in the space science domain rather than the Earth science domain and is relatively small, containing only 14 items. It meets few of the requirements defined here apart from providing and monitoring system feedback. Not all items are directly available on the site, making some easier to obtain than others. However, it does have some

smaller sized components that are more easily reused then larger, finished products, even if they are plug-ins designed to work with other software.

Astronomical Software and Documentation Service (ASDS)

The ASDS was funded by the NASA Astrophysics Data Program and run by the Space Science Telescope Institute. Its goal is "to allow astronomers and astronomical software developers to easily locate existing programs for their use and re-use, providing a uniform level of high-level documentation."

All users may view the site and its contents and follow links to the web pages hosting the assets. Site registration is not provided. Users may contribute assets by completing a web-based submission form at the simplest level or by installing a client package at the highest level. System feedback may be provided by e-mail to members of the ASDS team. There is no automatic notification system available. Assets may be located by full list, hierarchy, and search, though the search appeared to be broken. There is no apparent way to register usage of an asset or provide asset reviews, but some assets do have an "in our experience" section with comments from the team. Feedback monitoring is limited to system feedback. There are currently 78 entries listed in the catalog.

The site is easy to navigate and items are easy to find. Selecting one of the top-level hierarchy categories leads to a page that contains some description about the items in it; the full list is grouped by category but does not contain these descriptions. Ease of downloading depends on the external site hosting the asset. The assets are typically finished products, usually with source code, and some subroutines. There is a mix of free and non-free software. The domain is obviously astronomy. The site also contains a section dedicated to documentation such as manual/help pages for the listed software packages. Support is available by e-mail to the ASDS team and individual asset sites may provide additional means of support.

Requirement/Feature	Available at ASDS?
Domain	Astronomy
Type of Assets	Mainly packages, some subroutines, often with source
Register User	☆☆☆
Contribute/Update Assets	★★☆
Provide System Feedback	★★☆
Automatic Notifications	☆☆☆
Discovering Assets	List, Hierarchy, Search
Register Asset Usage	☆☆☆
Provide Asset Review	☆☆☆
Monitoring Feedback	★☆☆

Requirement/Feature	Available at ASDS?
Secure Log In / Registration	N/A
Catalog or Repository	Catalog
Operation Support	Six member team
Technology	HTML?

The domain of this system is astronomy, not Earth science, and it meets some of the requirements, but is lacking the ability to register users, provide automatic notifications, register asset usage, or provide asset reviews. While most of the packages are programs designed for the end user, some are smaller items such as subroutines that are more easily reused. Most of the assets are free, but there are some products listed that require a fee of some type (e.g., commercial or shareware). It appears that most asset pages have not been updated in over five years, so it is possible that the site is no longer actively maintained.

Other Sites Inspected

There are other software catalogs and repositories in addition to the ones reviewed here. Feedback from some members of the Earth science community indicated that they used a number of sites when looking for reusable assets. Some of them have been reviewed already (SourceForge, Freshmeat, Netlib, HDF-EOS Tools); some others are listed here. Web links for these sites are provided above.

The Direct Readout Laboratory hosted at NASA Goddard Space Flight Center is a repository containing 14 pieces of software and algorithms for use in this field, "the process of acquiring freely transmitted live satellite data". The site meets some of the requirements, but contains software for only one specific field within the Earth science domain. Therefore it is of limited use outside of that field.

The software repository hosted by the NASA Glenn Research Center contains 38 pieces of software for use in various engineering fields such as aerodynamics and fluid dynamics. The site meets some of the requirements, but does not focus on the Earth science domain, limiting its use to Earth science software developers.

ArcScripts is a repository of scripts designed to work with ESRI software which is for Earth science use, but apparently commercial. It is unclear whether or not the scripts in the repository can be used independently of the ESRI software. If not, their reuse value is limited since they are dependent on commercial software. The site appears to meet most of the requirements, but does not have a good way to provide asset reviews.

Wikipedia is a free encyclopedia. Entries often provide links to external sites with more information, so this site could be helpful in finding software assets in some cases. However, it is not a software repository or catalog and does not merit a full review. The same is true of Usenet newsgroups (e.g., available through http://groups.google.com/) – these are discussion groups and while useful information can be obtained through such a source, it cannot be reviewed with the repositories and catalogs here.

The Ruby Application Archive is a catalog of general Ruby software and SciRuby is a catalog of scientific Ruby software. The former tends to meet more of the requirements; the latter is a Wiki site which can be edited by anyone, and not specifically designed to function as a catalog. There seems to be little to no Earth science specific software at either location. Both sites are specific to one particular programming/scripting language and therefore are only useful if you are using that language. A software catalog/repository is more useful if it provides assets written in many different languages.

The Comprehensive Perl Archive Network is a general catalog of Perl software. The system seems to meet some of the requirements, but it covers a wide variety of areas, mainly in computing topics, so there is little Earth science software available. And as for the Ruby sites above, it is restricted to a single language, limiting its usefulness.

The FreeGIS Project has a software catalog of approximately 245 entries for use in the field of geographic information systems. It also contains a smaller number of other entries including documents and geo-data. The site meets some of the requirements, but covers one specific sub-field of Earth science, making it of limited use for any work outside of that field.

Review Summary

To facilitate comparison between the various systems, we present two summary tables with the results of our evaluation. The first table presents the NASA systems reviewed and the second table presents the non-NASA systems reviewed.

NASA Systems Reviewed

Requirement / Feature	Global Change Master Directory (GCMD)	GSFC Open Source Site	Ames Open Source Site	HDF-EOS Tools and Information Center	Computational Technologies Project	Earth Observing System Clearinghouse (ECHO)	Planetary Data Systems Software Download
Domain	Earth science	Earth and space science	General science	Earth science, HDF/HDF-EOS	Earth and space science	Earth science	Planetary astronomy
Type of Assets	Data sets, data services	Open source packages	Open source packages	Applications	Applications and source code	Metadata	Tools, binaries and source
Register User	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	★☆☆	☆☆☆
Contribute/Update Assets	★★☆	★☆☆	★☆☆	☆☆☆	☆☆☆	★★☆	☆☆☆
System Feedback	★★☆	★★☆	★★☆	☆☆☆	★★☆	★★☆	★★☆
Automatic Notifications	★★★	★☆☆	★☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Discovering Assets	Hierarchy, Search	List	List	List, Filter	Hierarchy	Search	List
Register Asset Usage	☆☆☆	★★★	★★★	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Provide Asset Review	☆☆☆	★☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Monitoring Feedback	★☆☆	★★☆	★☆☆	☆☆☆	★☆☆	★☆☆	★☆☆
Secure Log In / Registration	N/A	NO	NO	NO	N/A	YES	N/A
Catalog or Repository	Catalog	Both	Both	Repository	Catalog	Catalog	Both
Operation Support	Large	Small	Small	Inactive	Small	Available	Small
Technology	RSYNC, Zope, CVS, Linux, Java, and others	PHP	JavaServer Pages	Cold Fusion	HTML	XML (WSDL), SOAP, UDDI	Cold Fusion

Non-NASA Systems Reviewed

Requirement / Feature	Open Channel Foundation / COSMIC	SourceForge	Freshmeat	Scientific Applications on Linux	National Technology Transfer Center	National HPCC Software Exchange	Netlib	Savannah	Space Telescope Science Institute	Astronomical Software and Documentation Service
Domain	General	General	General	Scientific	Federal technologies (mostly NASA)	HPPC	Mathematics	General	Astronomy	Astronomy
Type of Assets	Applications and source code	Open source applications	Open source applications	Tools and packages with source code	Applications	Tools and end packages	Source codes	Tools and packages	Packages, source	Packages, source
Register User	★★★	★★★	★★☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	★★☆	☆☆☆	☆☆☆
Contribute/Update Assets	★★☆	★★★	★★☆	☆☆☆	★☆☆	☆☆☆	☆☆☆	★★☆	☆☆☆	★★☆
System Feedback	★☆☆	★☆☆	★★☆	☆☆☆	★★☆	☆☆☆	☆☆☆	★★☆	★★☆	★★☆
Automatic Notifications	★★☆	★★☆	★★★	☆☆☆	★★☆	☆☆☆	☆☆☆	★★☆	☆☆☆	☆☆☆
Discovering Assets	List, Hierarchy, Search	Hierarchy, Search	Hierarchy, Search	Hierarchy, Search (broken)	List, Hierarchy, Search	Hierarchy, Search	Hierarchy, Search	List, Search	List, Hierarchy	List, Hierarchy, Search
Register Asset Usage	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Provide Asset Review	☆☆☆	★☆☆	★★★	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Monitoring Feedback	★★☆	★★☆	★★☆	☆☆☆	★☆☆	☆☆☆	★☆☆	★★☆	★☆☆	★☆☆
Secure Log In / Registration	YES	YES	NO	N/A	N/A	N/A	N/A	YES	N/A	N/A
Catalog or Repository	Repository	Repository	Repository	Catalog	Both?	Catalog	Repository	Repository	Repository	Catalog
Operation Support	Medium	Large	Medium	Inactive	Uncertain	Inactive	Large	Large	Small	Medium
Technology	PHP, MySQL	PHP	XML-RPC	HTML, Java	ASP	Repository In a Box	HTML	Perl, PHP, MySQL	HTML	HTML

Conclusion

None of the sites evaluated fulfill the role of a software repository for the Earth science community. None of the catalogs and repositories that were evaluated sufficiently meet the requirements that are necessary to serve our community even though they may meet the needs of their targeted community well. In particular, the need to support multiple licensing mechanisms for reusable software assets and the need to restrict access to some assets to meet export control and NASA release requirements provide particular obstacles to using most of the system evaluated.

The Global Change Master Directory (GCMD) catalog is primarily a data catalog, offering information about Earth science data sets rather than software assets, it primarily serves as a catalog of metadata. The GCMD is not a repository, it provides links to other sites; it does not have the capability to host assets. In addition to serving as a catalog to data, it does list services (mainly software tools), but the number of services listed are limited in comparison to the number of data sets listed. The system also lacks the ability to provide asset reviews which is an important support feature. The GCMD does not target our intended audience. It focuses on providing information about data sets for data users, and little attention is paid to providing reusable software assets for software developers. Modifying the system to address software developers, as an afterthought, would leave them as the secondary market, possibly unable to have their needs met satisfactorily. The absence of repository and review capabilities would make it exceedingly difficult to adapt the GCMD to meet our needs. However, because of the GCMD's large user base and the close link between data and software, it is important that any new reuse enablement system provide the GCMD with a data feed of relevant content.

Similarly, the NASA Open Source Agreement web sites focus on one specific type of asset and license, preventing them from fully servicing the needs of the community of Earth science software developers. The NASA open source software sites at the Goddard Space Flight Center (GSFC) and the Ames Research Center are unable to satisfy the functional requirements. Neither site has an underlying catalog/repository system and assets are simply listed as links on a web page. Neither site can register users. In addition, providing asset reviews is either not available, as in the case of the Ames site, or available, as in the case of the GSFC site, but the reviews are not posted for the public. These two sites also have the common problem of primarily listing finished products rather than smaller software components.

Both sites host NASA open source software, but open source is just one way of licensing software for release and software developers should not have their choices restricted by license type. Also, software developers need to have access to software produced by other organizations, but listing non-NASA products would be inappropriate given the objectives of these sites. Related to both of these points, the sites seem to focus more on showing a sample of the open source software produced by NASA than actively promoting reuse. Given the restrictions on license and organization, this is almost inevitable. However, these restrictions on license and organization combined with the lack of catalog/repository capabilities render the sites unsuitable for meeting the needs of the software development community.

The Earth Observing System Clearinghouse (ECHO) is a data/service broker and has not been designed to act as a software repository or catalog for software developers. As such, ECHO does not appear to be a suitable candidate for hosting reusable software assets for the software development community.

General science/engineering sites are not domain-specific enough to meet the needs of the Earth science software development community. One such site is the Open Channel Foundation (OCF). OCF hosts the NASA COSMIC collection and contains a limited amount of Earth science material. The collection was initially hosted by the University of Georgia Research Foundation, but the relationship with NASA was terminated in 1998. The collection was moved from there to the OCF where it has become largely outdated. OCF meets some of our requirements, but also has some limitations. It does not have asset reviews (only forum capability), it is not focused on the Earth science domain, and it focuses on finished products rather than reusable software components. There are also concerns about using a third-party host including the ability to meet NASA release requirements and restricting access to NASA-only software. Therefore, the OCF is unsuitable for meeting the needs of the Earth science software development community.

General software repositories are even less domain-specific, making them incapable of meeting the needs of the Earth science software development community. SourceForge and Freshmeat are similar systems containing many assets from several fields. However, only a limited number of assets relevant to Earth science are available Both repositories make some attempt to target software developers by providing a location where they can store software for collaborative development, but many of the assets provided are finished products designed for end users, not software components that a developer could use in another project. Also, the focus of both of these repositories on open source software limits the users' choices by not making available software released under other types of licenses. The domain for each repository is essentially all-encompassing without any focus on a particular field or subject matter, making them unsuitable for meeting the specialized needs of our focused community.

Some of the sites community members mentioned using are useful within a very limited area, typically a specific programming or scripting language (such as Ruby or Perl) or one particular branch of Earth science (such as geographic information systems). While the target audience for these systems is mainly software developers, the domain is so specialized and not necessarily Earth science related that they cannot meet the needs of the general Earth science community. In addition, many of the sites tend to provide finished products which are generally not easily reused by software developers in their own projects. Community feedback has indicated that smaller sized components such as scripts are more desired for reuse than larger components such as executables. In general, most sites of all types we reviewed do not provide these types of assets.

Even though we concentrated on the Earth science community, we believe that the reuse needs of the space science community are similar. During this study we also did not come upon any catalogs/repositories that targeted the community of space science software developers well. It is likely that a similar trade study done for the space science community would come to the same conclusion. So there is a good possibility that a successful Earth science Reuse Enablement System could be reused later by the space science community with appropriate linkages between them for assets that are useful to both communities.

Having examined the sites specified by NASA HQ as well as additional NASA and non-NASA sites, we have determined that none of the existing systems perform the role of providing software developers in the Earth science community with the types of reusable assets they find most useful. While each system may meet the needs of its target audience, none have been designed to meet the reuse needs of the Earth science software development community. None of the existing systems satisfactorily meet all of the functional and non-functional requirements identified by the Working Group as being necessary to host a reuse enablement system for the Earth science community.

All of the above point towards establishing a new repository/catalog system, an Earth science Reuse Enablement System, as the best option for meeting the needs of the Earth science software development community. To meet the community's needs, a software repository/catalog system must do the following: meet the functional requirements, focus on the Earth science domain, target software developers as the primary audience, and provide the types of assets (including open source software and small-sized components) that are most desired by the community of Earth science software developers for reuse purposes.

Next Step

This study has focused on the evaluation of established catalog and repository systems. We have not evaluated in detail the technical options for establishing the Earth science Reuse Enablement System. A further architecture study would be needed to identify the most expeditious and cost-effective solution for such a system. The results of such a study might show that attempting to modify any of the existing systems to suit our purpose could dilute our requirements or unduly propose modifications that could compromise the functionality of an existing system. Our community's needs may not be met well since Earth science software developers would not be the primary audience of the system and members of this community may not have direct control of the system to make required changes and updates; lack of control when the system is hosted by a third-party may be a concern. In addition, major changes to a system that is functioning well for its designed purpose can negatively affect the system and create problems for the current users. Based on our existing knowledge of the available cataloging and repository technologies, we would expect that a Reuse Enablement System would be able to utilize existing technologies, including COTS and open source software, to fulfill a significant part of the required functionality. (For example, SourceForge Enterprise Edition, based on the collaborative software development tools that power SourceForge.net, is available on a commercial software basis from VA Software.) There is also a potential for some collaboration and data interchange between a new Reuse Enablement System and current catalogs and repositories such as the GCMD.

Appendix A – Software Reuse Questionnaire

The majority of the survey consisted of multiple choice questions where each listed option was ranked from 1 (not important at all) to 5 (very important). The following charts show the weighted average results for the top few responses to two of the questions.

Question 7 – How important were the following factors in preventing you from reusing software development artifacts developed outside your group?
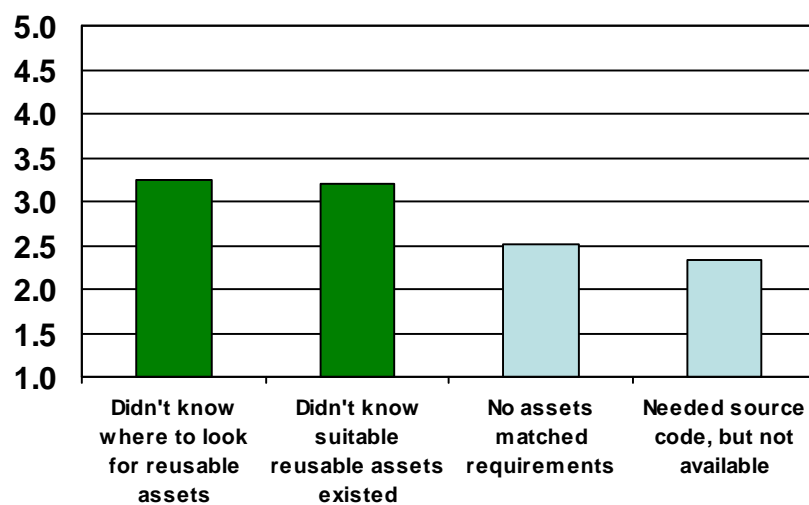
Question 47 – In your opinion, how important would the following factors be in helping increase the level of reuse within the Earth science community?
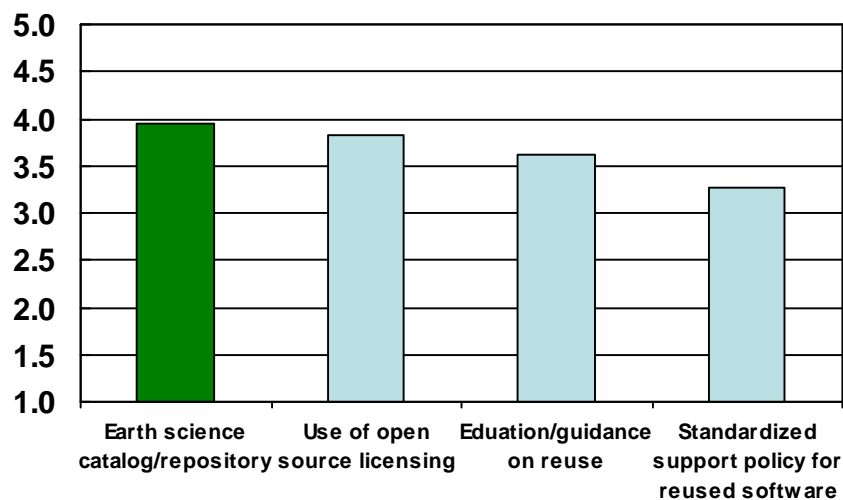
Appendix B – Enabling Systems Recommendation

The Software Reuse Working Group previously submitted a recommendation for a Reuse Enablement System to NASA HQ. This appendix contains the content of that recommendation and HQ's response to it.

- NASA should establish a system to facilitate the cataloging and distribution of reusable assets for the Earth science community

- NASA should establish an effective mechanism for dissemination of reusable assets within the Earth science community

- NASA should evaluate the technology options for the provision of a reuse enablement system including:

    - commercial reuse catalogs/repositories

    - open source reuse catalogs/repositories

    - use of existing publicly available catalogs/repositories

    - custom build of a community-specific catalog

- Based on the conclusions of the technology evaluation, NASA should implement a reuse enablement system

- NASA should develop guidelines and standards for the management and operation of a reuse enablement system

Impact for the Working Group

- The reuse working group will evaluate the technology options for the provision of a reuse enablement system

- The reuse working group will develop guidelines and standards for the management and operation of a reuse enablement system

- The reuse working group will develop a proposal for the implementation of a reuse enablement system based on the conclusions of the technology evaluation

- One additional FTE will be required for the balance of '05 fiscal year

Desired Decision

- HQ agreement to proceed with the evaluation of technology options and to provide funding for the evaluation

- HQ agreement in principle to the establishment of a reuse catalog subject to the findings of the evaluation

Headquarters' Response

- HQ thinks such a recommendation is premature and needs to await the results of a trade study concerning the establishment of a reuse catalog

Appendix C – General Requirements

For several months in 2004, the Software Reuse Working Group collaborated to define a set of requirements for a software Reuse Enablement System serving the Earth science community. This appendix contains a list of those requirements as previously submitted to NASA HQ.

General Requirements

- The system will facilitate the distribution and reuse of software development artifacts across the Earth and space science communities

- The reusable artifacts supported by the system will include software components and other digital artifacts used in the software development process

- The system will run on industry standard hardware and operating system

- The system will support remote access through standard Internet browsers

- The system will support the automated collection of system and asset usage metrics

- The system will provide error handling for all capabilities

- The system shall be flexible to support changes in NASA policy and strategy

Search Requirements

- The system will allow users to browse and look at system content without registering

- The system will allow users to discover (search for and find) assets of interest using multiple search mechanisms (e.g., keyword search or category search)

- The system will allow search results to be ordered in a number of ways (e.g., by category or rating)

User Registration

- The system will allow new users to register with the system and the user role defined by the registration will determine the user's access authority within the system

- Each user registration will require the approval of a system administrator

- The system will allow a user to update their user profile

- The system will allow registered users to provide system feedback

- The system will allow registered users to subscribe to system or asset events including events such as new versions, updates, and comments supplied by other users

Asset Usage

- The system will allow a Consumer to acquire an asset from the system repository

- The system will allow a Consumer to register usage of an asset, indicating active usage of the asset (this is different from downloading the asset)

- The system will provide a user forum for discussion and comments on assets

- The system will allow a Consumer to provide a rating and feedback on his/her experience with a particular asset

- The system will allow posting of requests for reusable assets that currently are not in the system

Asset Submission

- The system will allow a Provider to submit a new asset profile to the system

- The Provider may optionally upload artifacts associated with the asset

- Each asset submission will require the approval of a Content Manager before it can be accessed by other system users

- The system will allow a Provider to update the information about an asset and change the artifacts associated with the asset

- The system will allow Providers to subscribe to asset events including such events as comments and new requests pertaining to their contribution

Content Management

- The system will allow users to review feedback on assets and allow the Content Manager to remove feedback on assets (e.g., to make sure comments are on topic)

- The system will allow the Content Manager to review and approve asset submissions prior to them being made available to the community

- The system will allow the Content Manager to review the assets and remove those which are no longer relevant; this includes those that have poor reviews and/or no users

- The system will allow the Content Manager to review unsuccessful searches to capture consumer demand for assets that are not registered

System Administration

- The system will allow Administrators to monitor the general operating state of the system and perform designated routine tests to determine that the system is functioning properly

- The system will allow Administrators to manage user accounts and passwords

- The system will allow Administrators to monitor user feedback and use it to determine evolutionary needs of the system and other users

- The system will allow Administrators to generate reports including metrics

- The system will send notifications to subscribed users of system issues or events